BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G01N 37/00 — A2

(11) International Publication Number: **WO 99/63351**

(43) International Publication Date: 9 December 1999 (09.12.99)

(21) International Application Number: PCT/GB99/01742

(22) International Filing Date: 1 June 1999 (01.06.99)

(30) Priority Data:
9811656.9 29 May 1998 (29.05.98) GB

(71) Applicant *(for all designated States except US)*: OXFORD GLYCOSCIENCES (UK) LIMITED [GB/GB]; 10 The Quadrant, Abingdon Science Park, Abingdon OX14 3YS (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: PAREKH, Rajesh, Bhikhu [GB/GB]; Alchester House, Langford Lane, Near Wendlebury, Oxon OX6 0NS (GB). BRUCE, James, Alexander [GB/GB]; 14 Marlborough Crescent, Long Hanborough, Oxon OX8 8JP (GB). PHILP, Robin [GB/GB]; 27 Haywards Close, Wantage, Oxon OX12 7AT (GB).

(74) Agent: GILL JENNINGS & EVERY; Broadgate House, 7 Eldon Street, London EC2M 7LH (GB).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: GELS, METHODS AND APPARATUS FOR IDENTIFICATION AND CHARACTERIZATION OF BIOMOLECULES

(57) Abstract

The present invention is directed to efficient, computer–assisted methods and apparatus for identifying, selecting and characterizing biomolecules in a biological sample. According to the invention, a biological sample is first treated to isolate biomolecules of interest, and a two–dimensional array is then generated by separating the biomolecules present in a complex mixture. The invention provides a computer–generated digital profile representing the identity and relative abundance of a plurality of biomolecules detected in the two–dimensional array, thereby permitting computer–mediated comparison of profiles from multiple biological samples. This automatable technology for screening biological samples and comparing their profiles permits rapid and efficient identification of individual biomolecules whose presence, absence or altered expression is associated with a disease or condition of interest.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

GELS, METHODS AND APPARATUS FOR IDENTIFICATION AND CHARACTERIZATION OF BIOMOLECULES

1. INTRODUCTION

This invention relates to computer-assisted methods and apparatus for efficiently and systematically studying molecules that are present in biological samples and determining their role in health and disease. In particular, this invention relates to the emerging field of proteomics, which involves the systematic identification and characterization of proteins that are present in biological samples, including proteins that are glycosylated or that exhibit other post-translational modifications. The proteomics approach offers great advantages for identifying proteins that are useful for diagnosis, prognosis, or monitoring response to therapy and in identifying protein targets for the prevention and treatment of disease.

2. BACKGROUND OF THE INVENTION

Recent advances in molecular genetics have revealed the benefits of high-throughput sequencing techniques and systematic strategies for studying nucleic acids expressed in a given cell or tissue. These advances have highlighted the need for operator-independent computer-mediated methods for identifying and selecting subsets or individual molecules from complex mixtures of proteins, oligosaccharides and other biomolecules and isolating such selected biomolecules for further analysis.

Strategies for target-driven drug discovery and rational drug design require identifying key cellular components, such as proteins, that are causally related to disease processes and the use of such components as targets for therapeutic intervention. However, present methods of analyzing biomolecules such as proteins are time consuming and expensive, and suffer from inefficiencies in detection, imaging, purification and analysis.

Though the genomics approach has advanced our understanding of the genetic basis of biological processes, it has significant limitations. First, the functions of products encoded by identified genes -- and especially by partial cDNA sequences -- are frequently unknown. Second, information about post-translational modifications of a protein can rarely be deduced from a knowledge of its gene sequence, and it is now apparent that a large proportion of proteins undergo post-translational modifications (such as glycosylation and phosphorylation) that can

profoundly influence their biochemical properties. Third, protein expression is often subject to post-translational control, so that the cellular level of an mRNA does not necessarily correlate with the expression level of its gene product. Fourth, automated strategies for random sequencing of nucleic acids involve the analysis of large numbers of nucleic acid molecules prior to determining which, if any, show indicia of clinical or scientific significance.

For these reasons, there is a need to supplement genomic data by studying the patterns of protein and carbohydrate expression, and of post-translational modification generally, in a biological or disease process through direct analysis of proteins, oligosaccharides and other biomolecules. However, technical constraints have heretofore impeded the rapid, cost-effective, reproducible, systematic analysis of proteins and other biomolecules present in biological samples.

3. SUMMARY OF THE INVENTION

The present invention is directed to efficient, computer-assisted methods and apparatus for identifying, selecting and characterizing biomolecules in a biological sample. According to the invention, a biological sample is first treated to isolate biomolecules of interest, and a two-dimensional array is then generated by separating the biomolecules present in a complex mixture. The invention provides a computer-generated digital profile representing the identity and relative abundance of a plurality of biomolecules detected in the two-dimensional array, thereby permitting computer-mediated comparison of profiles from multiple biological samples. This automatable technology for screening biological samples and comparing their profiles permits rapid and efficient identification of individual biomolecules whose presence, absence or altered expression is associated with a disease or condition of interest. Such biomolecules are useful as therapeutic agents, as targets for therapeutic intervention, and as markers for diagnosis, prognosis, and evaluating response to treatment. This technology also permits rapid and efficient identification of sets of biomolecules whose pattern of expression is associated with a disease or condition of interest; such sets of biomolecules provide constellations of markers for diagnosis, prognosis, and evaluating response to treatment.

The high throughput, automatable methods and apparatus of the present invention further permit operator-independent selection of individual separated

biomolecules (or subsets of separated biomolecules) according to pre-ordained criteria, without any requirement for knowledge of sequence information or other structural characteristics of the biomolecules. This in turn provides automated, operator-independent isolation and parallel characterization of a plurality of selected biomolecules detected in a biological sample. Thus, the present invention advantageously permits automated selection of biomolecules prior to sequencing or structural characterization.

4. DETAILED DESCRIPTION OF THE INVENTION

The present invention provides methods for rapidly and efficiently identifying and characterizing bio-molecules, for example proteins, in a biological sample. In a first step, a biological sample is treated to isolate the biomolecules of interest prior to separating the biomolecules for characterisation. Purification is performed with a view either to selectively enrich certain desirable biomolecules, e.g. proteins, from within the sample or to selectively deplete the sample of certain undesirable biomolecules. For example, if glycosylated proteins only are of interest, such glycoproteins may be selectively isolated from a sample using lectin-affinity chromatography or lectin affinity precipitation. Such enrichment can both enhance and simplify the subsequent protein separation and analysis. Any proteins or group of proteins carrying a structural determinant for which an antibody or other specific purification reagent is available may be so extracted, e.g. tyrosine phospho-proteins by using an anti-phosphotyrosine antibody. Conversely, a sample may be depleted of specific proteins, again using protein-specific affinity methods. For example, albumin may be removed from body fluids using an anti-albumin antibody, and immunoglobulins may be removed using protein A or protein G (preferably immobilised) and haptoglobin and transferrin can be similarly removed. It is clear that a sample may be selectively depleted (or enriched) for more than one protein by using protein-specific reagents serially or in combination. Such enrichment/depletion can often have a beneficial effect during analysis, by concentrating proteins of interest/removing proteins that interfere with or may, for example by their predominance, limit the analysis of proteins of interest.

In either case of enrichment or depletion, the end result is to provide a sample containing the biomolecules of interest in a more isolated or "pure" form with respect to the original biological sample. A preferred method for achieving this is affinity chromatography e.g. Fast Protein Liquid Chromatography (FPLC).

After purification, the sample containing the biomolecules is subjected to two successive separation steps. In the first separation step, the biomolecules are separated according to one physical or chemical property so as to generate a one-dimensional array containing the biomolecules; for example, proteins are separated by isoelectric focusing along a first axis. In the second separation step, the biomolecules in this one-dimensional array are separated according to a second physical or chemical characteristic so as to generate a two-dimensional array of separated biomolecules; for example, proteins separated by isoelectric focusing are subjected to SDS-PAGE along a second axis perpendicular to the first axis. The separated biomolecules are stably maintained in the two-dimensional array for subsequent imaging. The stable two-dimensional array can be stored or archived for an extended period (*e.g.* months or years) and selected biomolecules can be retrieved from the array at any desired time, based on automated computer analysis of the data derived from imaging.

The two-dimensional array is imaged with a detector to generate a computer-readable output that contains a set of x,y coordinates and a signal value for each detected biomolecule. If desired, the computer-readable output can be displayed to a human operator -- before or after computer-mediated analysis -- as a computer-generated image on a screen or on any suitable medium. Computer-mediated analysis of the computer-readable output is performed, resulting in a computer-readable profile that represents, for a plurality of detected biomolecules, the relative abundance of each such biomolecule and its attributes as deduced from its x,y coordinates in the two-dimensional array. For example, a profile derived from imaging a gel containing proteins separated by isoelectric focusing followed by SDS-PAGE represents the isoelectric point (pI), apparent molecular weight (MW) and relative abundance of a plurality of detected proteins.

The computer-readable profiles of the present invention are suitable for computer-mediated analysis to identify one or more biomolecules that satisfy specified criteria. In one embodiment, a first set of profiles is compared with a second set of profiles to identify biomolecules that are represented in all the profiles of the first set (or in a first percentage of the profiles of the first set) and are absent from the profiles of the second set (or are absent from a second percentage of the profiles of the second set, where the first and second percentages can be independently specified). In other embodiments, sets of

profiles are compared to identify biomolecules that are present at a designated higher level of expression in a specified percentage of profiles of one sample set than in a specified percentage of profiles of another sample set, or to identify biomolecules whose post-translational processing differs from one sample set to another.

One or more biomolecules so identified are selected for isolation. In one embodiment, this selection is made automatically by a computer, in accordance with pre-ordained programmed criteria, without further human intervention. In another embodiment, a human operator reviews the results of the computer-mediated analysis and then enters a selection into a computer. For isolation of each selected biomolecule, a computer generates machine-readable instructions that direct a robotic device (a) to remove one or more portions of the two-dimensional array that contain the selected biomolecule and (b) to deliver the removed portions to one or more suitable vessels for further characterization. For example, a selected protein can be analyzed to determine its full or partial amino acid sequence, to detect and characterize any associated oligosaccharide moieties, and to study other aspects of post-translational processing, *e.g.* phosphorylation, myristylation and the like. The invention advantageously permits automated parallel processing of biomolecules removed from the two-dimensional array, thereby facilitating rapid and efficient characterization of a plurality of selected biomolecules. Figure 1 of WO-A-9823950 presents a flowchart illustrating processing of a sample.

The present invention is useful for identifying and analyzing proteins, but is more generally applicable to the identification and analysis of any biomolecule. As used herein, the term "biomolecule" refers to any organic molecule that is present in a biological sample, and includes peptides, polypeptides, proteins, oligosaccharides, lipids, steroids, prostaglandins, prostacyclines, and nucleic acids (including DNA and RNA). As used herein, the term "protein" includes glycosylated and unglycosylated proteins. These and other terms and procedures are as defined in WO-A-9823950; see the sections headed "Biological samples", Analysis of proteins" (twice), "Analysis of oligosaccharides", "Computer analysis of the detector output", "Computer generation and analysis of profiles", "Removal of selected portions of a supported gel", and "Processing removed portions of the gel".

In connection with the first "Analysis of proteins" section, a purification step is now introduced. A wide variety of purification techniques may be used in this first step. For example, purification may occur by the use of Fast Protein Liquid Chromatography (FPLC), ion exchange chromatography or affinity chromatography. Preferably FPLC is used, comprising one or more affinity columns containing affinity chromatography media which binds selectively the biomolecules of interest.

5. EXAMPLE: PROTEINS FROM SERUM AND SYNOVIAL FLUID OF PATIENTS WITH RHEUMATOID ARTHRITIS

Proteins in serum and synovial fluid from patients with rheumatoid arthritis (RA) were purified by FPLC, separated by isoelectric focusing followed by SDS-PAGE and compared.

5.1. FPLC Purification

Selected proteins which are desired to be specifically removed from the sample prior to proteome analysis e.g. albumin, haptoglobin and transferrin present in serum were removed by a FPLC purification step. This was achieved by passing the biological sample through a series of Hi-trap affinity chromatography columns each comprising immobilised antibodies specific for a particular protein. The specific proteins bind to the column and the eluate is collected and concentrated by centrifugal ultrafiltration.

5.2. Isoelectric Focusing

For isoelectric focusing (IEF), each sample (after FPLC treatment) was applied to an Immobiline® DryStrip Kit (Pharmacia BioTech), following the procedure described in the manufacturer's instructions, *see* Instructions for Immobiline® DryStrip Kit, Pharmacia, # 18-1038-63, Edition AB (incorporated herein by reference in its entirety), with optional modifications as described by Sanchez et al. 1997, Electrophoresis 18: 324-327 (incorporated herein by reference in its entirety).

In certain cases, in order to increase the resolution in a particular pH range or to load a larger quantity of a target protein onto the gel, a narrow-range "zoom gel" having a pH range of 2 pH units or less was used, according to the method described in Westermeier, 1993, Electrophoresis in Practice (VCH, Weinheim, Germany), pp. 197-209 (which is incorporated herein by reference in its entirety).

5.3. Gel Equilibration and SDS-PAGE

IEF gels were prepared for SDS-PAGE by equilibration in a SDS buffer system according to a two step procedure comprising initial reduction of the disulfide bonds, followed by alkylation of the free thiol groups, as described by Sanchez et al., *id.* Thereafter, SDS-PAGE was carried out according to Hochstrasser et al., 1988, Analytical Biochemistry 173: 412-423 (incorporated herein by reference in its entirety), with modifications as specified below.

5.4. Preparation of supported gels

Covalent attachment of SDS-PAGE gels to a glass support was achieved by applying a 0.4% solution of γ-methacryl-oxypropyltrimethoxysilane in ethanol to the glass plate ("the bottom plate") to which the gel was to be attached. Excess reagent was removed by washing with water, and the bottom plate was allowed to dry. At this stage, both as identification for the gel, and as a marker to identify the coated face of the plate, an adhesive bar-code was attached to the bottom plate in a position such that it would not come into contact with the gel matrix.

An opposing glass plate ("the top plate") was treated with RepelSilane (Pharmacia Biotech) to minimize gel attachment. After applying the reagent, the top plate was heated by applying a flow of heated air (*e.g.* from a hot air gun) to the surface of the plate. Excess reagent was again removed by water washing, and the top plate was allowed to dry.

The dried plates were assembled into a casting box with a capacity of 13 gel sandwiches. Several casting boxes can be assembled in parallel to cast more gels under the same conditions. The top and bottom plates of each sandwich were spaced by means of 1mm thick spacers. The sandwiches were interleaved with acetate sheets to facilitate separation of the sandwiches after gel polymerization. Casting was then carried out according to Hochstrasser et al., op. cit.

5.5. SDS-PAGE

The gel strips from the IEF step were applied to the top of the poured SDS-PAGE gel and electrophoresis begun. In order to ensure even cooling of the gel during the electrophoresis run, a system was designed essentially as described by Amess et al,. 1995, Electrophoresis 16: 1255-1267 (incorporated herein by reference in its entirety). Even, efficient cooling is desirable in order to minimize

thermal fluctuations during electrophoresis and hence to maintain the consistency of migration of the proteins. Electrophoresis was carried out until the tracking dye reached the bottom edge of the gel. The gels were then removed immediately for staining.

5.6. Staining

The top plate of the gel cassette was carefully removed, leaving the gel bonded to the bottom plate. The bottom plate with its attached gel was then placed into a staining apparatus, which has the capacity to accommodate 12 gels. The gels were completely immersed overnight in fixative solution, comprising 40% (v/v) ethanol, 10% (v/v) acetic acid, 50% (v/v) water. The fixative was then drained from the tank, and the gels were primed by immersion in 7.5% (v/v) acetic acid, 0.05% (w/v) SDS for 30 mins. The priming solution was then drained, and the gels were stained by complete immersion in the dye solution for 4 hours. A stock solution of fluorescent dye was prepared by diluting Sypro Red (Molecular Bioprobes, Inc., Eugene, Oregon), according to the manufacturer's instructions. The diluted solution was filtered under vacuum though a 0.4 μm filter.

In order to achieve a continuous, even circulation of the various solutions over all 12 gels, solutions were introduced into the tank via a distribution bar, extending along the bottom of the tank across its entire width and provided with holes that allow the solution to flow evenly over each of the gels.

5.7. Imaging of the gel

A computer-readable output was produced by imaging the fluorescently stained gels with a Storm scanner (Molecular Dynamics, Sunnyvale, California) according to the manufacturer's instructions, (*see* Storm User's Guide, 1995, Version 4.0, Part No. 149-355, incorporated herein by reference in its entirety) with modifications as described below. Since the gel was rigidly bonded to a glass plate, the gel was held in contact with the scanner bed during imaging. To avoid interference patterns arising from non-uniform contact between the gel and the scanner bed, a film of water was introduced under the gel, taking care to avoid air pockets. Moreover, the gel was placed in a frame provided with two fluorescent buttons that were imaged together with the gel to provide reference points (designated M1 and M2) for determining the x,y coordinates of other features detected in the gel. A matched frame was provided on a robotic gel excisor in order to preserve accurate alignment of the gel. After imaging, the gels were

sealed in polyethylene bags containing a small volume of staining solution, and then stored at 4°C.

The output from the scanner was first processed using MELANIE® to autodetect the registration points, M1 and M2; to autocrop the images (*i.e.*, to eliminate signals originating from areas of the scanned image lying outside the boundaries of the gel, *e.g.* the reference frame); to filter out artifacts due to dust; to detect and quantify features; and to create image files in GIF format. Features were detected by a computer-mediated comparison of potential protein spots with the background to select areas of the gel associated with a signal that exceeded a given threshold representing background staining.

A second program was used for interactive editing of the features detected and to match duplicate gels for each sample. First, images were evaluated to reject images which had gross abnormalities, or were of too low a loading or overall image intensity, or were of too poor a resolution, or where duplicates were too dissimilar. If one image of a duplicate was rejected then the other image belonging to the duplicate was also rejected regardless of image quality. Samples that were rejected were scheduled for repeat analysis.

Landmark identification was used to correct for any variability in the running of the gel. This process involves the identification of certain proteins which are expected to be found in any given biological sample. As these common proteins exhibit identical isoelectric points and molecular weight from sample to sample, they can be used as standards to correct for any possible gel variation or distortion. The pI and molecular weight values for the landmarks in the reference gel were determined by co-running a sample with *E. coli* proteins which had previously been calibrated with respect to known protein in human plasma. Features which were considered to be artifacts, mainly at the edges of the gel image and particularly those due to the sample application point and the dye-front, were removed. Duplicate gels were then aligned via the landmarks and a matching process performed so as to pair identical spots on the duplicate gels. This provided increased assurance that subsequently measured isoelectric points and molecular weights were accurate, as paired spots demonstrated the reproducibility of the separation. The corrected gel, in addition to being used for subsequent analysis, was printed out for visual inspection.

Generation of the image was followed by computer measurement of the x,y coordinates of each protein, which were correlated with particular isoelectric points and molecular weights by reference to the known landmark proteins or standards. A measurement of the intensity of each protein spot was taken and stored. Each protein spot was assigned an identification code and matched to a spot on a master gel, *i.e.*, a reference gel which contained most or all of the protein spots seen in each type of sample and was used as a template to which the protein spots of the other samples were matched. This step allowed for the identification of putative correlate spots across many different gels. The data collected during collection of the original biological sample, as described in section 5.1, were reunited with the gel data, thereby permitting the analysis of computer selected cross-sections of the samples based on information such as age or clinical outcome.

The end result of this aspect of the analysis was the generation of a digital profile which contained, for each identified spot: 1) a unique arbitrary identification code, 2) the x,y coordinates, 3) the isoelectric point, 4) the molecular weight, 5) the signal value, 6) the standard deviation for each of the preceding measurements, and 7) a pointer to the MCI of the spot on the master gel to which this spot was matched. By virtue of the LIMS, this profile was traceable to the actual stored gel from which it was generated, so that proteins identified by computer analysis of gel profile databases could be retrieved. The LIMS also permitted the profile to be traced back to the original sample or patient.

5.8. Digital Analysis of the Gel

Once the profile was generated, analysis was directed toward the selection of interesting proteins.

The protein features in the individual images from the paired serum and synovial fluid samples were compared electronically. Molecular identity of any one feature across the set of images is defined in this analysis as identity of position in the 2-D separation. Quantitative measurement of the abundance of an individual feature in an individual image was based on normalized fluorescence intensity measured for that feature in that image. Those proteins whose abundance differed between the sets of serum and synovial fluid samples were revealed by electronic comparison of all detected features in all relevant images.

5.9. Recovery and analysis of selected proteins

Differentially expressed proteins were robotically excised and processed to generate tryptic peptides; partial amino acid sequences of these peptides were determined by mass spectroscopy, using de novo sequencing.

5.10 Results

These initial experiments identified 12 proteins that were present at higher levels in human RA synovial fluid than in matched serum samples, and 9 proteins that were present at lower levels in human RA synovial fluid than in matched serum samples. Partial amino acid sequences were determined for each of these differentially expressed proteins. Computer analysis of public databases revealed that 16 of these partially sequenced proteins were known in the art and that 5 were not described in any public database examined.

References herein to US Patent Application No. 08/877,605 apply also to the International Patent Application No. PCT/GB98/01486. In addition, reference may be made to WO-A-9801749, for a discussion of techniques for the enhanced separation of species, e.g. in a gel.

CLAIMS

1. A computer-assisted method for selecting and directing the isolation of one or more biomolecules present in a two-dimensional array, comprising:

a purification step, wherein a plurality of biomolecules of interest are substantially isolated from a first biological sample;

a first separation step, wherein said biomolecules are separated according to a first physical or chemical property to form a one-dimensional array of biomolecules;

a second separation step, wherein said one-dimensional array of biomolecules is separated according to a second physical or chemical property to form said two-dimensional array;

imaging said two-dimensional array or a replica thereof to generate a computer-readable output comprising, for each of a plurality of biomolecules detected in said two-dimensional array, a pair of x,y coordinates and a signal value;

processing said output in at least one computer to select one or more of said detected biomolecules in accordance with previously ordained or operator-specified criteria; and optionally

generating machine-readable instructions that direct a robotic device to isolate at least one of said selected biomolecules from said two-dimensional array.

2. The method according to claim 1, further comprising:

isolating at least one of said selected biomolecules from said two-dimensional array by means of said robotic device in accordance with said machine-readable instructions.

3. The method according to claim 1 or claim 2, in which said biomolecules are oligosaccharides.

4. The method according to claim 1 or claim 2, in which said biomolecules are proteins.

5. The method according to claim 4, in which said proteins are glycoproteins.

6. The method according to any preceding claim, in which said two-dimensional array is contained in a polyacrylamide gel.

7. The method according to claim 6, in which said biomolecules have been separated by isoelectric focusing, followed by electrophoresis in the presence f sodium dodecyl sulfate.

8. The method according to claim 6 or claim 7, in which said polyacrylamide gel is bonded to a generally planar solid support such that the gel has two-dimensional spatial stability, and the support is substantially non-interfering with respect to detection of a detectable label carried by the proteins.

9. The method according to claim 8, in which said polyacrylamide gel is covalently bonded to said solid support.

10. The method according to claim 8 or claim 9, in which said detectable label is a fluorescent label.

11. The method according to any of claims 8 to 10, in which said solid support is glass.

12. The method according to any preceding claim, wherein the purification step is carried out using FPLC.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G01N 27/447 — A3

(11) International Publication Number: WO 99/63351

(43) International Publication Date: 9 December 1999 (09.12.99)

(21) International Application Number: PCT/GB99/01742

(22) International Filing Date: 1 June 1999 (01.06.99)

(30) Priority Data:
9811656.9 29 May 1998 (29.05.98) GB

(71) Applicant *(for all designated States except US)*: OXFORD GLYCOSCIENCES (UK) LIMITED [GB/GB]; 10 The Quadrant, Abingdon Science Park, Abingdon OX14 3YS (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: PAREKH, Rajesh, Bhikhu [GB/GB]; Alchester House, Langford Lane, Near Wendlebury, Oxon OX6 0NS (GB). BRUCE, James, Alexander [GB/GB]; 14 Marlborough Crescent, Long Hanborough, Oxon OX8 8JP (GB). PHILP, Robin [GB/GB]; 27 Haywards Close, Wantage, Oxon OX12 7AT (GB).

(74) Agent: GILL JENNINGS & EVERY; Broadgate House, 7 Eldon Street, London EC2M 7LH (GB).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report:
27 January 2000 (27.01.00)

(54) Title: METHOD FOR IDENTIFICATION AND CHARACTERISATION OF BIOMOLECULES

(57) Abstract

The present invention is directed to efficient, computer–assisted methods and apparatus for identifying, selecting and characterizing biomolecules in a biological sample. According to the invention, a biological sample is first treated to isolate biomolecules of interest, and a two–dimensional array is then generated by separating the biomolecules present in a complex mixture. The invention provides a computer–generated digital profile representing the identity and relative abundance of a plurality of biomolecules detected in the two–dimensional array, thereby permitting computer–mediated comparison of profiles from multiple biological samples. This automatable technology for screening biological samples and comparing their profiles permits rapid and efficient identification of individual biomolecules whose presence, absence or altered expression is associated with a disease or condition of interest.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

INTERNATIONAL SEARCH REPORT

Interna al Application No
PCT/GB 99/01742

A. CLASSIFICATION OF SUBJECT MATTER

G01N27/447

According to International Patent Classification (IPC) or to both national classification and IPC6

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5073963 A (SAMMONS et al.) 17 December 1991. --	1
A	WO 98/19271 A (MOSE LARSEN, P.) 07 May 1998. --	1
A	WO 96/39625 A (BELTRONICS INC.) 12 December 1996. ----	1

☐ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 September 1999	22 11. 1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.O. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	NARDAI e.h.

Form PCT/ISA/210 (second sheet) (July 1992)

ANHANG

zum internationalen Recherchenbericht über die internationale Patentanmeldung Nr.

ANNEX

to the International Search Report to the International Patent Application No.

ANNEXE

au rapport de recherche international relatif à la demande de brevet international n°

PCT/GB 99/01742 SAE 236790

In diesem Anhang sind die Mitglieder der Patentfamilien der im obengenannten internationalen Recherchenbericht angeführten Patentdokumente angegeben. Diese Angaben dienen nur zur Unterrichtung und erfolgen ohne Gewähr.

This Annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The Office is in no way liable for these particulars which are given merely for the purpose of information.

La présénte annexe indique les membres de la famille de brevets relatifs aux documents de brevets cités dans le rapport de recherche international visée ci-dessus. Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsibilité de l'Office.

Im Recherchenbericht angeführtes Patentdokument Patent document cited in search report Document de brevet cité dans le rapport de recherche	Datum der Veröffentlichung Publication date Date de publication	Mitglied(er) der Patentfamilie Patent family member(s) Membre(s) de la famille de brevets	Datum der Veröffentlichung Publication date Date de publication
US A 5073963	17-12-1991	AU A1 80019/91 WO A1 9119274	31-12-1991 12-12-1991
WO A1 9819271	07-05-1998	AU A1 45693/97 AU A1 54070/98 EP A1 932882 EP A2 934409 WO A2 9820124 WO A3 9820124 AU A1 41325/97 EP A1 925554 WO A1 9811508	22-05-1998 29-05-1998 04-08-1999 11-08-1999 14-05-1998 08-10-1998 02-04-1998 30-06-1999 19-03-1998
WO A1 9639625	12-12-1996	AU A1 57756/96 EP A1 871873 US A 5865975	24-12-1996 21-10-1998 02-02-1999